



RESEARCH ARTICLE

REVISED Stage 2 Registered Report: Variation in neurodevelopmental outcomes in children with sex chromosome trisomies: testing the double hit hypothesis

[version 4; peer review: 2 approved]

Dianne F. Newbury ¹, Nuala H. Simpson ², Paul A. Thompson ², Dorothy V. M. Bishop ²

¹Department of Biological and Medical Sciences, Oxford Brookes University, Oxford, OX3 0BP, UK

²Department of Experimental Psychology, University of Oxford, Oxford, OX2 6GG, UK

V4 First published: 19 Jul 2018, 3:85
<https://doi.org/10.12688/wellcomeopenres.14677.1>
 Second version: 10 Oct 2018, 3:85
<https://doi.org/10.12688/wellcomeopenres.14677.2>
 Third version: 07 Sep 2020, 3:85
<https://doi.org/10.12688/wellcomeopenres.14677.3>
 Latest published: 01 Jun 2021, 3:85
<https://doi.org/10.12688/wellcomeopenres.14677.4>

Abstract

Background: The presence of an extra sex chromosome is associated with an increased rate of neurodevelopmental difficulties involving language. The 'double hit' hypothesis proposes that the adverse impact of the extra sex chromosome is amplified when genes that are expressed from the sex chromosomes interact with autosomal variants that usually have only mild effects. We predicted that the impact of an additional sex chromosome on neurodevelopment would depend on common autosomal variants involved in synaptic functions.

Methods: We analysed data from 130 children with sex chromosome trisomies (SCTs: 42 girls with trisomy X, 43 boys with Klinefelter syndrome, and 45 boys with XYY). Two comparison groups were formed from 370 children from a twin study. Three indicators of phenotype were: (i) Standard score on a test of nonword repetition; (ii) A language factor score derived from a test battery; (iii) A general scale of neurodevelopmental challenges based on all available information. Preselected regions of two genes, *CNTNAP2* and *NRXN1*, were tested for association with neurodevelopmental outcomes using Generalised Structural Component Analysis.

Results: There was wide phenotypic variation in the SCT group, as well as overall impairment on all three phenotypic measures. There was no association of phenotype with *CNTNAP2* or *NRXN1* variants in either the SCT group or the comparison groups. Supplementary

Open Peer Review

Reviewer Status

Invited Reviewers

	1	2
version 4 (revision) 01 Jun 2021	 report	
version 3 (update) 07 Sep 2020	 report	 report
version 2 (revision) 10 Oct 2018		
version 1 19 Jul 2018	 report	 report

1. **Beate St Pourcain** , Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
2. **David H. Skuse** , UCL Great Ormond Street Institute of Child Health, London, UK

analyses found no indication of any impact of trisomy type on the results, and exploratory analyses of individual SNPs confirmed the lack of association.

Conclusions: We cannot rule out that a double hit may be implicated in the phenotypic variability in children with SCTs, but our analysis does not find any support for the idea that common variants in *CNTNAP2* or *NRXN1* are associated with the severity of language and neurodevelopmental impairments that often accompany an extra X or Y chromosome.

Stage 1 report: <http://dx.doi.org/10.12688/wellcomeopenres.13828.2>

Keywords

sex chromosome trisomy, language, autism spectrum disorder, neuroligin, synapse, X chromosome, Y chromosome

Any reports and responses or comments on the article can be found at the end of the article.

Associated Study Protocol

Newbury DF, Simpson NH, Thompson PA and Bishop DVM » Stage 1 Registered Report: Variation in neurodevelopmental outcomes in children with sex chromosome trisomies: protocol for a test of the double hit hypothesis, *F1000Research* 2018, **3**:10 (<https://doi.org/10.12688/wellcomeopenres.13828.2>)

Corresponding author: Dorothy V. M. Bishop (dorothy.bishop@psy.ox.ac.uk)

Author roles: **Newbury DF:** Conceptualization, Data Curation, Formal Analysis, Methodology, Project Administration, Supervision, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Simpson NH:** Conceptualization, Data Curation, Investigation, Methodology, Project Administration, Validation, Writing – Review & Editing; **Thompson PA:** Data Curation, Formal Analysis, Methodology, Project Administration, Software, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Bishop DVM:** Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Software, Supervision, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This work was funded by Wellcome Trust Programme Grants [082498], and European Research Council Advanced Grant [694189].

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2021 Newbury DF *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Newbury DF, Simpson NH, Thompson PA and Bishop DVM. **Stage 2 Registered Report: Variation in neurodevelopmental outcomes in children with sex chromosome trisomies: testing the double hit hypothesis [version 4; peer review: 2 approved]** Wellcome Open Research 2021, **3**:85 <https://doi.org/10.12688/wellcomeopenres.14677.4>

First published: 19 Jul 2018, **3**:85 <https://doi.org/10.12688/wellcomeopenres.14677.1>

REVISED Amendments from Version 3

In this version we have provided further analytic details requested by reviewers, and reworded in a few places for clarity. No new analyses have been added, though we have corrected omission of one page of results in Appendix 9, as noted by reviewer DHS.

Some confusion had been caused in version 3 by the addition of novel analyses under the heading of 'Exploratory analyses', given that we already had 'additional exploratory analyses' in version 2. We have revised headings to clarify which material was new in version 3.

Any further responses from the reviewers can be found at the end of the article

Introduction

Developmental language disorder (DLD), a condition in which there are unexplained and persistent difficulties with language acquisition, affects around 7% of children (Norbury *et al.*, 2016). Family studies show that DLD runs in families (Bishop, 2008), yet it has proved hard to identify any genetic or environmental factors that substantially increase risk. One reason is that DLD appears to be a complex multifactorial disorder where influences of individual genetic variants (alleles) are typically of small effect and may interact with other genetic factors and with the environment. Indeed, the ways in which disorders pattern in families suggest that common genetic variants that confer risk of language disorder may lead to an autistic phenotype when they occur with other genetic risk factors (Bishop, 2010). Thus the specific phenotype can depend on the constellation of genetic variants, rather than there being separate risk factors for DLD and autism spectrum disorder (ASD).

Rather than recruiting increasingly large numbers to try to find reliable associations between language disorders and genetic variants in genome-wide studies, one way forward is to study rare disorders that have a large impact on the phenotype, which may point to functional pathways involved in more common forms of disorder. One instance of a striking association between a genetic condition and language disorder in children of normal intelligence is provided by the sex chromosome trisomies (SCTs), each of which affects 1–1.5 per 1000 children (Nielsen & Wohlert, 1991). In the 1960s, research was initiated to investigate neurodevelopmental outcomes of children with SCT detected on neonatal screening. A systematic review of these studies showed that in all three trisomies there were high rates of speech and language impairment, motor problems, and educational difficulties, despite IQ being within normal limits in most cases (Leggett *et al.*, 2010). Furthermore, studies of samples who have developmental language disorder of unknown cause find an increased prevalence of sex chromosome trisomies (Simpson *et al.*, 2014).

In a study of children with sex chromosome trisomies identified on prenatal screening, Bishop *et al.* (2011) found that 7 of 30 (24%) girls with karyotype 47,XXX, 9 of 19 (47%) boys with 47, XXY and 15 of 21 (71%) boys with 47,XYY had a history of speech and language-therapy, compared with rates of

4% in sisters and 18% in brothers. Furthermore, this same study found that 2 of 19 (11%) boys with 47,XXY, and 4 of 21 (20%) boys with 47,XYY had received a diagnosis of ASD, compared with an estimated national prevalence rate of 0.2% in girls and 0.6% in boys. In addition, many children with SCTs who were not diagnosed with ASD had evidence of communication difficulties on parental report, including pragmatic (autistic-like) problems, in all three karyotypes. More recent research has provided further evidence of a link with autism as well as other neurodevelopmental disorders in boys with a sex chromosome trisomy (Ross *et al.*, 2012).

The impact of a trisomy is influenced by distinctive characteristics of the sex chromosomes. In most cases, the phenotypic effects of SCTs are much less severe than the impact of an autosomal trisomy: Down syndrome (trisomy 21) usually leads to intellectual disability, and most other trisomies are lethal. Viable trisomies usually involve small chromosomes with a low gene count (for example the Y chromosome), where the effects associated with altered gene dosage are less severe. An exception to this rule is the X chromosome. The X chromosome has a relatively high gene count, but the impact of a duplication is relatively mild because mechanisms of inactivation have evolved, such that in typical females, only one copy is active, and in effect, both males and females have one set of functional genes from this chromosome. In trisomies that involve the X chromosome, two copies are inactivated, largely negating the presence of additional genetic material. There are, however, exceptions to this rule, with between 12–20% genes escaping inactivation to some extent: These include genes in the pseudo-autosomal region, and other genes that have homologues on the Y chromosome (Carrel & Brown, 2017).

The fact that there is an increase in problems affecting speech, language and communication in all three sex chromosome trisomies suggests there is an adverse impact of an additional copy of a gene that is expressed and has homologous forms on the X and Y chromosomes. *Neurologin-4 (NLGN4)* is a strong candidate for such a gene, for several reasons (Bishop *et al.*, 2011). First, *NLGN4X*, located on Xp22, at least partly escapes inactivation (Berletch *et al.*, 2011). Second there is a homologous gene, *NLGN4Y* on the Y chromosome at Yq11.2. Third, *neurologins* are expressed in brain, as well as other tested tissues (Abrahams & Geschwind, 2010; Jamain *et al.*, 2003). Fourth, as reviewed by Cao & Tabuchi (2017), mutations of *NLGN4* have been linked to ASD (Jamain *et al.*, 2003; Laumonier *et al.*, 2004; Lawson-Yuen *et al.*, 2008; Marshall *et al.*, 2008; Pampanos *et al.*, 2009; Talebizadeh *et al.*, 2006; Yan *et al.*, 2008) – although this finding is inconsistent and other studies have not found autism in those with mutations of *NLGN4* (Chocholska *et al.*, 2006; Macarov *et al.*, 2007), or have failed to find abnormalities of *NLGN4* in those with autism (Blasi *et al.*, 2006; Gauthier *et al.*, 2005; Liu *et al.*, 2013; Vincent *et al.*, 2004; Yanagi *et al.*, 2012). Fifth, *neurologins* are post-synaptic transmembrane proteins that mediate development of functional synapses between neurons and are in the same functional network as neuroligins (Craig & Kang, 2007), which have also been implicated in both DLD and ASD (Vernes *et al.*, 2008). Jamain *et al.* (2003) proposed that a defect in *NLGN4* may abolish formation or function of synapses involved in

communication. Note that these authors also implicated another X-chromosome neuroligin, *NLGN3*, in autism, but this is located at Xq13, where one copy would be inactivated, and there is no homologue on the Y-chromosome. Therefore, unlike *NLGN4*, *NLGN3* would not be over-expressed in those with an extra X or Y chromosome.

For the reasons described above, we may hypothesise that an extra copy of *NLGN4* could be implicated in neurodevelopmental problems. However, we also need to explain within-karyotype variation. Although there is a substantial increase in rates of speech, language and social communication problems in children with SCTs, the additional chromosome does not cause language impairment or ASD in a deterministic fashion. A minority of children have no evidence of developmental difficulties, a minority are severely affected with disabilities extending across many domains, and most have mild to moderate impairments (Linden & Bender, 2002).

The wide variation in outcomes suggests that the extra gene dosage could act as a multiplier of other risk factors, which interact with the sex chromosome genes in a dosage-dependent manner and so only assume importance in the subset of individuals who have other genetic or environmental risk factors (Bishop & Scerif, 2011). This explanation is consistent with rodent research comparing the effect of a *NLGN3* mutation between different strains of mouse, suggesting the impact is dependent on the genetic background (Jaramillo *et al.*, 2018). It also is compatible with evidence from studies of mutations in *NLGN4* in humans, which found that the same mutation may be associated with different phenotypes within one family (Jamain *et al.*, 2003; Laumonnier *et al.*, 2004; Lawson-Yuen *et al.*, 2008; Yan *et al.*, 2005). As well as autism, *NLGN4* associations have been described with intellectual disability, language disorder and Tourette syndrome (Lawson-Yuen *et al.*, 2008; Yan *et al.*, 2005).

Hypothesis

Our pre-planned analysis was designed to test the ‘double hit’ hypothesis

The ‘double hit’ hypothesis: Neuroligins act as a multiplier of effects of neuroligins

The notion of a ‘double hit’ aetiology has been proposed previously to account for cases where a microdeletion is inconsistently associated with neurodevelopmental disorder (Girirajan *et al.*, 2010; Newbury *et al.*, 2013): the idea is that a severe phenotype may be seen when there are two copy number variants or mutations, each of which may be relatively innocuous on its own. Here, we extend that idea to argue that the effect of altered *neuroligin* gene dosage may depend on the genetic background provided by autosomes (Bishop & Scerif, 2011). In this regard, it is of particular interest to note that neuroligin proteins form part of the same functional network as a group of presynaptic transmembrane proteins, known as neuroligins; their interactions play a key role in synaptogenesis (Hussain & Sheng, 2005). *CNTNAP2* encodes a member of the neuroligin superfamily whose polymorphisms have been associated with common forms of language impairment (Graham & Fisher, 2015), though the effect size is relatively

small (Vernes *et al.*, 2008). The role of the *CNTNAP2* protein in developing brain is not fully understood, and it is likely to play multiple roles at different time-points. While early functional studies of the *CNTNAP2* protein indicated that it localises to nodes of Ranvier in axonal membranes, it is now recognised to have key functions at the synapse (Lu *et al.*, 2016; Zweier *et al.*, 2009). This raises the possibility that a *CNTNAP2* gene variant that has a modest effect in individuals of normal karyotype might have a much larger impact in the context of overexpression of a *neuroligin*. This hypothesis predicts that presence of an additional sex chromosome will amplify the impact of common genetic variants that have two characteristics: (a) they have been associated with DLD or ASD, and (b) they are in the same functional network as *neuroligins*. Figure 1 is a schematic showing two genes of interest to our current study, *CNTNAPs* and Neuroligins, interacting with neuroligins in the synaptic cleft.

Methods

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study (Simmons *et al.*, 2012).

Power analysis and impact of ascertainment bias

We aimed to recruit sufficient children with trisomies to detect an effect size of $d = 0.5$ for each copy of a given genetic variant on a phenotype, equivalent to a standardized regression

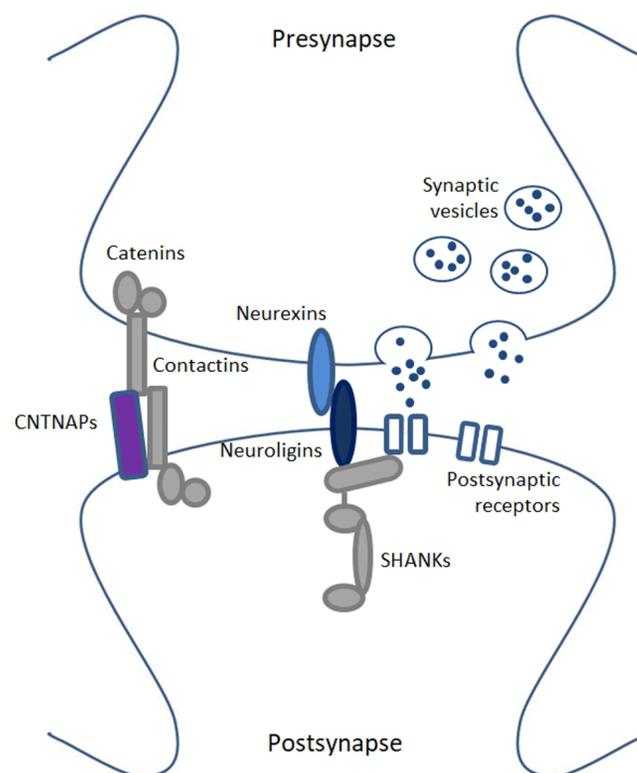


Figure 1. Neuroligins (such as *NRXN1*), neuroligins (such as *NLGN4*) and contactin-associated proteins (such as *CNTNAP2*) all form part of the synaptic scaffolding system.

slope of 0.25. The anticipated effect size is hard to judge, but the average impact of a sex chromosome trisomy on verbal IQ is more than one SD from the general population mean (Leggett *et al.*, 2010), suggesting that if the trisomy acts as a multiplier of effects of autosomal variants, this effect could be large. When testing variants with a prior association with disorder, we can make a directional prediction. We aimed to recruit 150 children with trisomies, which would have given 94% power to detect a slope of 0.25 on one-tailed test. However, we recruited only 140 children and had missing data on some variables, so numbers, and consequently power, are lower than this. In addition, we have to take into account that the sample is not representative of children with sex chromosome trisomies, because around 50% had the trisomy discovered in childhood when developmental difficulties were being investigated (see below). We devised a simulation to check the impact of these factors on power (see Appendix 1). This showed that a combination of $N = 130$ with 50% postnatally identified (and presumably biased) cases with mean phenotype score 0.9 SD below the group average (computed from a language factor score), reduced power to 87% on one-tailed test.

Participants

Sex chromosome trisomies: After excluding children with missing or inadequate DNA, participants included 42 girls with trisomy X, 43 boys with Klinefelter syndrome, and 45 boys with XYY. These were combined in a single group of 130 children for analysis, but are shown broken down by trisomy and background in Figure 2. Cases were recruited from National Health Service Clinical Genetics Centres, from two support groups (Unique: the Rare Chromosome Support Group, and the Klinefelter Syndrome Association), or from self-referral via

advertisements on the OSCCI website and our Facebook page. A criterion for inclusion was that the child was aware of their trisomy status. In a previous study (Bishop *et al.*, 2011) we noted that levels of impairment tended to be lower in cases where the trisomy was discovered on prenatal screening than in those identified later in childhood. We therefore asked parents specifically about the reason for genetic testing; for 59 children aneuploidy only came to light because of behavioural or developmental problems. Note that this means that data from this sample should not be used to estimate prevalence of neurodevelopmental disorders in sex chromosome trisomies.

Comparison group: Comparison data came from a sample of children aged from 6 years 0 months to 11 years 11 months who had completed the same test battery, who were taking part in a twin study of language and laterality (Wilson & Bishop, 2018a), and whose first language at home was English. Although twinning is a risk factor for early language delay, this effect appears to wash out with age, and by school age, genetic factors play a major role in the aetiology of language disorder (Bishop, 2006; Rice *et al.*, 2018). In this sample, we aimed for an over-representation of twin pairs in which one or both twins had language or literacy problems that might be indicative of DLD. This was coded on the basis of parental response on a telephone interview: any mention of language delay, history of speech and language therapy, current language problems or dyslexia was coded as 'parental concern'. We aimed to recruit 180 pairs selected on the basis of having language or literacy problems (60 MZ, 60 DZ opposite sex and 60 DZ same sex), and 60 unselected pairs (20 of each type); we fell short of this goal as seen in Figure 3. For the current analysis, we grouped together all twins, regardless of zygosity and parental concern, and then divided

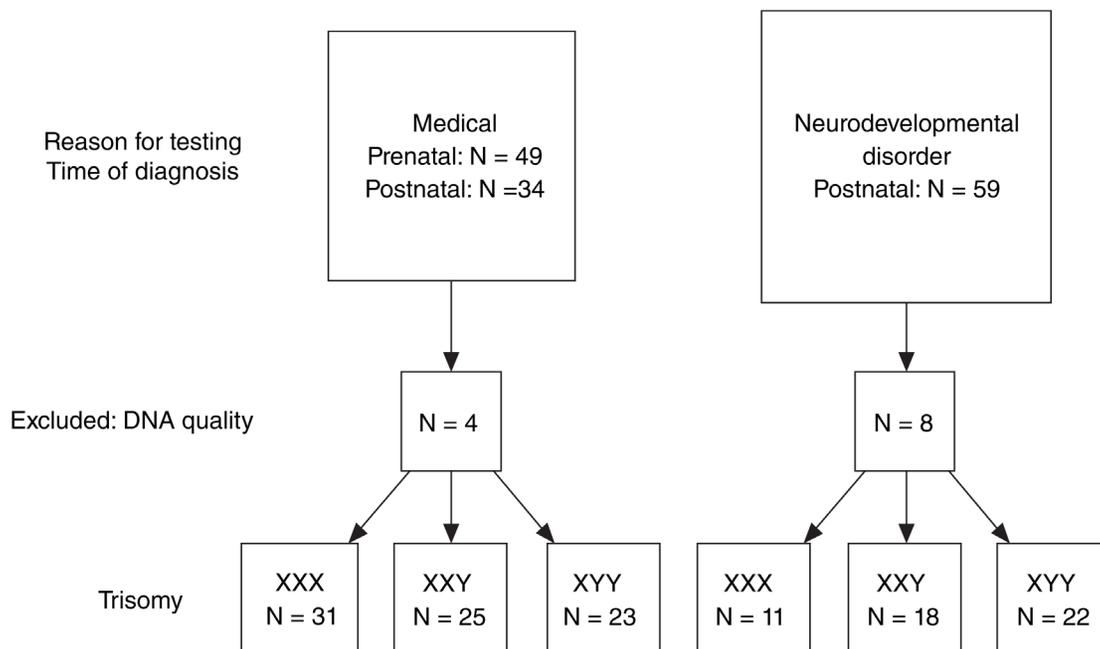


Figure 2. Flowchart showing characteristics of children recruited to sex chromosome trisomies (SCT) group.

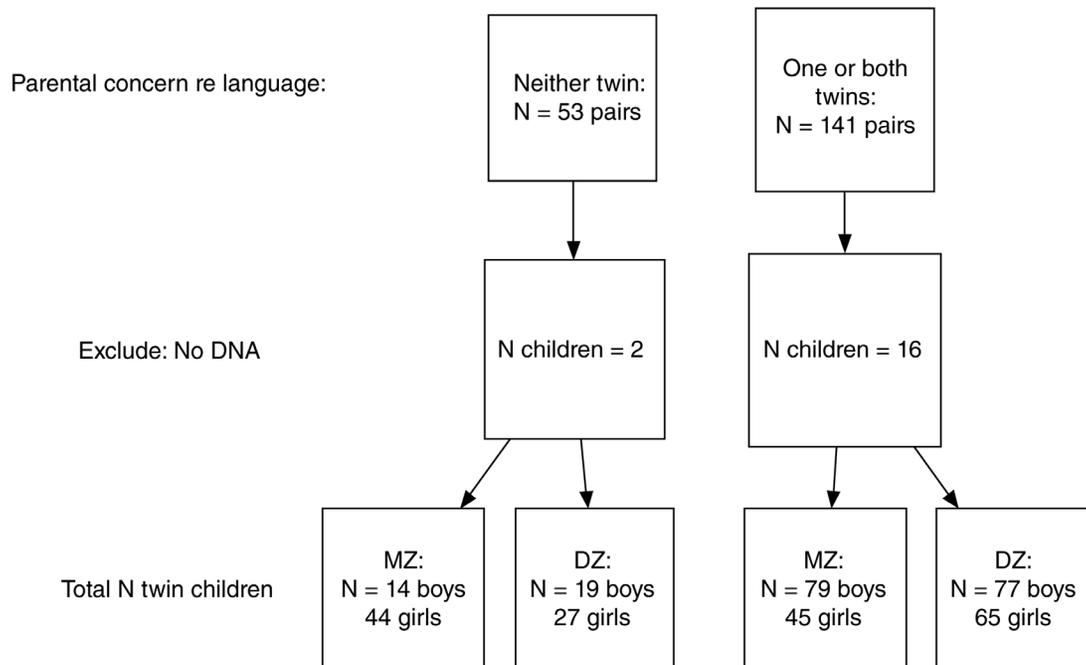


Figure 3. Flowchart showing characteristics of children recruited to comparison groups.

them into two subsamples by selecting one twin from each pair at random, after excluding 18 cases with missing or insufficient DNA. This means we can replicate the analysis for twins with a diploid (typical) karyotype. Note that this replication sample is not independent, as the genotype for the MZ twins is the same in the two subsamples, and is related for DZ twins.

Information about zygosity, gender and parental concern is shown for information, but was not used in the analysis. Because twins are not independent, the final sample was divided into two subgroups of 184 and 186 children respectively, each containing one member from each pair, selected at random. (Ns not equal because some twins had missing DNA from just one member of the pair).

Some twin children had evidence of autism spectrum disorder (N = 15) or intellectual disability (N = 3), and twelve failed a hearing screen on the day of testing, although none of them had any known sensorineural hearing loss. For the current study, because we were interested in a broader phenotype than pure DLD, these cases were retained in the sample.

Test battery

Psychiatric evaluation. In an initial telephone interview, parents were asked about the child's medical and educational history, including a question about whether anyone had diagnosed the child with a neurodevelopmental disorder such as ASD, developmental language disorder (DLD) or specific language impairment, dyslexia or dyspraxia. In addition, one or both parents

were asked to complete the online [Development and Wellbeing Assessment \(DAWBA\)](#) (Goodman *et al.*, 2000) in their own time. 84 parents of SCT cases and 133 parents of twins complied with this request. The DAWBA gives information on likelihood of the child meeting criteria for a range of psychiatric diagnoses; a final diagnosis is made by a trained rater who assimilates all the information and evaluates it against DSM5 criteria (American Psychiatric Association, 2013).

Language, literacy and cognitive assessments. All children were seen at home or in a quiet space in their school for a neurocognitive assessment, using the battery of language and nonverbal ability tests shown in [Table 1](#). Hearing was screened in left and right ears using a DSP Pure Tone Audiometer (Micro Audiometric Corporation). The child was familiarised with the task of raising their hand on hearing a tone using 40 dB (HL) tones. They were then tested with 25 dB pure tones at frequencies of 500, 1000, 2000 and 4000 Hz. Louder tones were presented in 5 dB steps to establish a threshold at any frequency where a 25 dB tone was not detected. Children with an average threshold greater than 30 dB in the better ear were categorized as failing the screen. The battery also included tests of literacy: the Picture and Digit naming tests from the Phonological Assessment Battery (Frederickson *et al.*, 1997), the Test of Word Reading Efficiency (Torgesen *et al.*, 1999) and the Neale Analysis of Reading Ability -2 (Neale, 1999), but these are not included in the current analysis as there was much missing data from the youngest children. In addition, handedness and language laterality were assessed. Results from laterality assessments were unremarkable and are not considered further here (Wilson & Bishop, 2018a; Wilson & Bishop, 2018b).

Table 1. Assessment battery.

Instrument	Measure
Woodcock Johnson III Tests of Cognitive Abilities (Woodcock <i>et al.</i> , 2007)	Verbal Comprehension
NEPSY: A Developmental Neuropsychological Assessment (Korkman <i>et al.</i> , 1998)	Repetition of Nonsense Words
	Oromotor Sequences
	Sentence Repetition
Wechsler Abbreviated Scale of Intelligence (WASI) (Wechsler, 1999)	Vocabulary
	Block Design
	Matrices
Parental questionnaires	
The Children's Communication Checklist-2 (CCC-2) (Bishop, 2003)	
The Social Responsiveness Scale (SRS) (Constantino, 2005)	

For the NEPSY tests, norms extend only to age 12 yr 11 months, and so we used extrapolated scores, as documented in [Appendix 2](#).

Phenotypes

We considered three quantitative phenotypes ranging from a specific measure of a heritable language skill, through a more general language measure, to a measure that potentially indexes a wide range of neurodevelopmental problems:

A) Nonword repetition, which is regarded as a measure of phonological short-term memory. This was singled out as an individual measure because it has previously been identified in twin studies as a good marker of heritable language problems (Bishop *et al.*, 1996) and has also been associated with genetic variants linked to language/literacy in the *CNTNAP2*, *CMIP*, *ATP2C2*, *KIAA0319*, and *DCDC2* genes (Carrion-Castillo *et al.*, 2017; Marino *et al.*, 2012; Newbury *et al.*, 2009; Newbury *et al.*, 2011; Scerri *et al.*, 2011; Vernes *et al.*, 2008). In the current study, we used scaled scores from Repetition of Nonsense Words from the NEPSY (Korkman *et al.*, 1998).

B) A general language factor derived from the four other language tests (Verbal Comprehension, Oromotor Sequences, Sentence Repetition and Vocabulary). As documented in [Appendix 3](#), the decision to combine these measures into a single language factor was made after exploring the factor structure of the available phenotypic measures, with the goal of obtaining a reliable indicator of overall language function.

C) A global measure of burden of neurodevelopmental problems extending beyond language, including autistic features. This was developed on an *ad hoc* basis, using all available information from parental report (see [Appendix 4](#)).

DNA collection and analysis

Oragene kits (OG-500, DNA Genotek Inc, Ontario, Canada) were used to collect saliva for DNA analysis from children with

SCTs and their parents and available twin pairs. DNA extraction was performed using an ethanol precipitation protocol as detailed in the standard protocol (DNA genotek). All extracted DNA was genotyped on the Infinium 'Global Screening Array-24 (v1)', which includes 692,824 SNPs including rare and common variations. Data were processed in the Illumina [BeadStudio/GenomeStudio](#) software (v. 2.03) and all SNPs with a GenTrain (quality) score of < 0.5 were excluded at this stage. All genotypes were further filtered using [PLINK](#) software v1.07 (Purcell *et al.*, 2007); as recommended by Anderson *et al.* (2010), samples with a genotype success rate below 95% or a heterozygosity rate ± 2 SD from the mean were removed, as were SNPs with a Hardy-Weinberg equilibrium $P < 0.000001$ or a minor allele frequency of less than 1%. Identity data within families and twin-pairs were used to exclude samples with unexpected gender or relationships. SNPs that showed an inheritance error rate > 1% or skewed missing rates between genotype plates were also excluded. Control data (CEU, YRI, CHB, JPT, Hapmap release #3) were employed through a principal component analysis within [Eigenstrat](#) (Price *et al.*, 2006) to identify individuals with divergent ancestry. Sixteen individuals (6 twin pairs and 4 SCT cases) were identified as having African ancestry and 21 individuals (6 twin pairs and nine SCT family members) were identified as having Asian ancestry. Any SNPs that showed a significant association with non-European ancestry ($P < 0.0001$) were excluded. The final genome-wide dataset consisted of 500 individuals (370 twins, divided into two subgroups, and 130 independent SCT cases) and 451,093 autosomal SNPs with a genotyping rate of 99.78%.

Procedure

Ethical approval was obtained for the study in 2011 from the Berkshire NHS Research Ethics Committee (reference 11/SC/0096), and data collection started in August of that year, finishing in October 2016. Information sheets, consent forms and ethics approval documents are available on [Open Science Framework](#). Families who had expressed interest in the study

were interviewed by telephone to assess whether the child met inclusion criteria, and if so, an appointment was made to see the child at home or at school, depending on parental preference. Families were widely dispersed around the UK, including Northern Ireland, Scotland, Wales and England. During the course of recruitment a total of eight research assistants as well as the senior author were involved in assessing children. The assessment was conducted in a single session lasting between 2–3 hours per child, with breaks where needed.

Analysis plan

Study data were analysed using [R software](#) ([R Core Team, 2016](#)), with the main database managed using [REDCap electronic data capture tools](#) hosted at the University of Oxford ([Harris et al., 2009](#)).

Potentially, there is a very large number of genotypes and phenotypes that could be analysed to test our hypothesis, as well as different ways of creating subgroups. This consideration, coupled with the small sample size, makes it important to control adequately for multiple testing to guard against type I error ([Grabitz et al., 2018](#)). For this reason, we stored phenotype and genotype data separately and specified an analysis plan in detail, as reported in our stage 1 registered report ([Newbury et al., 2018](#)). The analysis of genotype-phenotype associations was conducted after this plan had been registered and peer-reviewed.

Subgroups

In our main pre-specified analysis we treated all three trisomies together. This is because the double hit hypothesis postulates a common mechanism that would apply regardless of karyotype. We specified that if we found an association between genotype and phenotype, we would carry out exploratory analyses to consider whether this is moderated by karyotype. This would allow us indirectly to test a prediction by [Skuse \(2018\)](#) that there is more variable expressivity of *NLGN4X* than *NLGN4Y*. If so, one might expect a more severe impact of a double hit in the XYY group, which would be reflected in the epistatic interactions with *NLGN*, and should lead to lower phenotypic variability in XYY compared to the other karyotypes. Note, however, that the ascertainment bias in the sample is problematic for making cross-karyotype comparisons, and the focus would have to be just on those who were not diagnosed because of neurodevelopmental problems (see [Figure 2](#)). This is a small sample and so there would be a high risk of missing a true effect (type II error).

Prioritising genotypes for analysis

We conducted a series of literature searches to prioritise autosomal genes for analysis, focusing on genes that had an association with childhood speech and language disorders and that were relevant for synaptic function (see [Appendix 5](#)). This led us to select two candidates; *CNTNAP2* and *NRXN1*. Both of these genes are large (>1 MB) and included over 100 SNPs from the genotyping array. In order to avoid false positives with our small sample size, we chose to focus our analysis on regions that have previously been associated with neurodevelopmental disorder, analysing all genotyped SNPs (after quality control

steps described above (see “DNA collection and analysis”) within these selected regions.

In *CNTNAP2* ([NM_014141](#)), we focused on a region spanning exons 13–14 (chr7:147,514,390–147,612,852 (hg19)). This region includes a cluster of 9 SNPs previously associated with language disorder ([Vernes et al., 2008](#); [Whitehouse et al., 2011](#)). We had direct genotype data for 22 SNPs across this region. In addition, we used imputation to obtain genotypes for SNPs rs2710102 and rs7794745. These were the first SNPs reported to be associated with ASD, and represent the two main SNPs used in the majority of association studies in neurodevelopmental disorders ([Alarcón et al., 2008](#); [Arking et al., 2008](#)). These two SNPs were not directly genotyped on the Illumina arrays and were therefore imputed for all individuals. Imputation was performed on the [Michigan Imputation Server](#), an online server which generates phased and imputed genotypes using high-density reference panels. Variant Call Files (VCF) were uploaded for 15,936 SNPs genotyped on chromosome 7. Genotypes were phased within Eagle and imputed by Minimac against the Human Reference Panel hrc.r1.1.2016, which includes 64,940 haplotypes of predominantly European ancestry. In total, genotypes were generated for 2,289,829 SNPs across chromosome 7, 513,970 of which had quality scores 0.9. The two SNPs of interest, rs2710102 and rs7794745 had quality scores of 0.9938 and 0.94127 respectively.

The second candidate is *NRXN1* ([NM_004801](#)). Although this gene met our criteria of being relevant for both synaptic function and neurodevelopmental phenotypes, the studies showing this link involved deletions rather than common variants ([Ching et al., 2010](#)). A recent analysis of clinical microarray data showed that deletions in exons near the 5' end of *NRXN1* were specifically implicated in neurodevelopmental disorders ([Lowther et al., 2017](#)). Accordingly, we focused on 23 SNPs in this region of the gene chr2:51,141,501–51,280,121 (hg19). These SNPs covered exons 1–4 plus 20 Kb upstream (5') of the gene as this region includes important regulatory sequences. Details of the SNPs included in the analysis are shown in [Appendices 6 and 7](#).

The SNPs within the chosen regions were filtered for minor allele frequency and Hardy Weinberg Equilibrium (as outlined in DNA collection and analysis) but were not pruned for linkage disequilibrium. Previous simulations indicate that the Generalized Structured Component Analysis (GSCA) method is not greatly affected by linkage disequilibrium (see [Romdhani et al., 2014](#)). Across the *CNTNAP2* region, six pairwise combinations of SNPs had $R^2 > 0.8$. Across the *NRXN1* region, 8 pairwise combinations of SNPs had $R^2 > 0.8$. Details of the SNPs included in the analysis and a table of correlations between SNPs are shown in [Appendix 8](#).

Statistical methods

CNTNAP2 and *NRXN1* genes from the trisomy sample were analysed for association with a latent variable based on the three phenotypes using a structural equation modelling (SEM) approach adapted for genetic analysis ([Romdhani et al., 2015](#)). The model specification for our analysis is shown in [Figure 4](#).

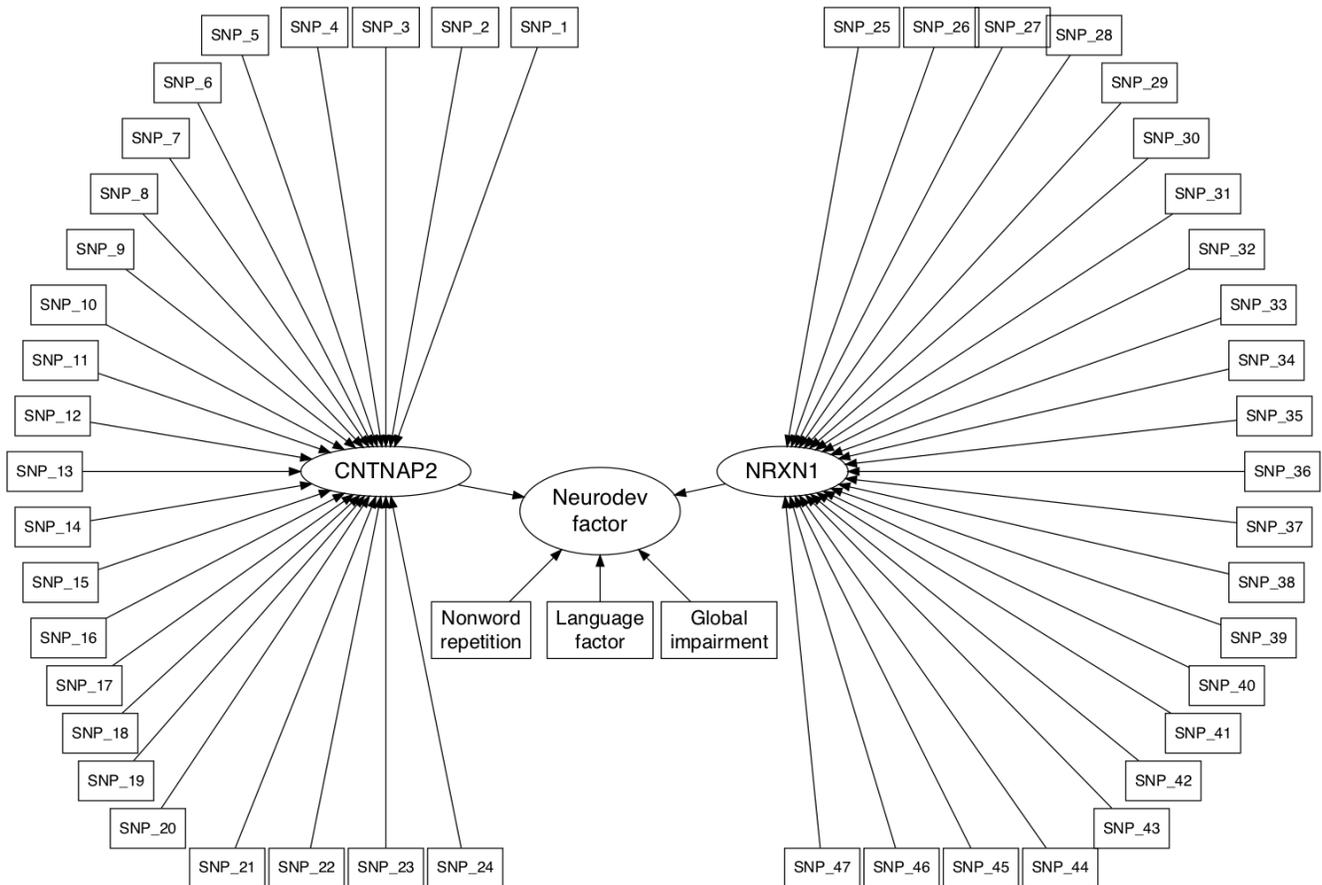


Figure 4. Structural equation diagram for analysis.

The GSCA analysis estimates the path from each gene to the Neurodev factor, with significance calculated by permutation analysis.

Romdhani *et al.* (2015) used the GSCA developed by Hwang & Takane (2004). This method uses component-based path modelling rather than traditional covariance-based SEM, allowing adequate model fit to be achieved when using smaller samples (Chin & Newstead, 1999; Tenenhaus *et al.*, 2005). The measurement models in the SEM framework are not typical regression format using latent factors; instead they are fitted using alternating least squares to estimate the weights and parameters, which is similar to principal components analysis. The advantage of this approach is that it does not attempt to fit the whole covariance matrix for observed and latent variables, but rather fits a separate measurement model for the contribution of observed variables to each latent factor, as well as a covariance model for the latent factors. Hence, we do not estimate the contribution of individual SNPs in each gene to the phenotype; rather, their influence is represented via the weighted sum. Similarly, the latent phenotypic factor (termed Neurodev factor in Figure 4) is a weighted sum of the three measures of the phenotype. We estimated the significance of one direct pathway from the *CNTNAP2* gene to the latent phenotype, and one from *NRXN1*

to the latent phenotype. This method thus gives a single estimate of the overall impact of SNPs in a region of the phenotype.

We conducted simulations that indicated that this method is feasible with the number of SNPs and phenotypes in our sample (see Appendix 1): the permutation method, used by this approach to effectively quantify the test statistic distribution, generates p-values independently for each path, and a correction is required to take this into account. Because the evidence of association of common variants was stronger for *CNTNAP2* than for *NRXN1*, we used a sequential approach to setting a significance level (alpha), using a critical p-value of .05 to test the pathway from *CNTNAP2* to the Neuro factor, and .025 for the pathway from *NRXN1* to the Neuro factor.

In addition, we conducted the same analyses with children from the two comparison samples.

We predicted that one or both paths from *CNTNAP2* and *NRXN1* to the Neurodev factor would indicate significant association in the sex chromosome trisomy sample. We further predicted that any associations in the comparison samples will be similar in direction, but smaller in size and may not reach statistical significance.

Further exploratory analyses (update added to version 3 of paper)

Gene-based analyses. Following the publication of the GSCA analyses described above, we performed gene-based analyses of the 22 candidate genes described in our analysis plan (Appendix 5, Figure S1). These methods simultaneously consider genetic variation across a gene unit in a single individual, incorporating both risk and protective effects. These effects are considered in a cumulative model in which a higher number of risk variants is associated with poorer outcome at the behavioural level. Such analyses can handle large numbers of variants and allow us to include variants that are individually rare. Note, however that gene-based approaches are still limited to consideration of single genes and often involve many parameters, increasing the potential for false positives. This approach is an exploratory analysis that may lead to the identification of a functional targeted subset of variants for further investigation. If rare variants were found to have an effect, then these could be followed up using GSCA analysis (Thompson *et al.*, 2020) as performed for *CNTNAP2* and *NRXN1* here.

Quantitative association analysis was performed for common and rare variants within the 18 autosomal genes and four X chromosome genes identified in a literature search as having an association with childhood speech and language disorders and relevant synaptic function (Figure S1). Through an imputation step, these analyses included all known variants across the 22 candidate genes. We tested for association with each of the three outcome measures of interest; nonword repetition, language factor and neurodevelopmental factor.

Genotypes were imputed per chromosome using the Michigan Imputation server (Das *et al.*, 2016) in the sex chromosome trisomy (SCT) samples and Twin sample sets. Genotypes from the Infinium 'Global Screening Array-24 (v1)' were extracted for the chromosomes containing the candidate genes and prepared for imputation in the following way. The Non-SNPs, (indels and multiallelic variants), A/T and G/C allele SNPs were excluded. Inconsistent reference alleles were also identified. These were then 'forced' so that the reference allele changed to the other allele or 'flipped' to the other strand, in PLINK v1.9. The vcf was then uploaded to the Michigan Imputation server using the following options; Human Reference panel - HRCr1.1 2016, Phasing - Eagle v2.3, Population - EUR, Mode - Quality Control and Imputation. For chromosome X all the options were the same apart from the Phasing - ShapeIT v2.r790 (unphased) as recommended by the Michigan Imputation Server. This reference panel consists of 64 976 haplotypes of predominantly European ancestry for 32 488 individuals across over 39 million variants genome-wide.

For each candidate gene, as annotated in hg19, the imputed SNPs were extracted for that region +/- 10kb either side. SNPs that had an R2 (quality score) <0.7 in the Michigan Imputation server '.info' file were excluded, as were SNPs that appeared twice in the output (multi-allelic variants). The

allele frequencies were checked and monomorphic SNPs were excluded. Hardy Weinberg Equilibrium was checked in vcftools and any SNPs with $p < 10^{-6}$ were excluded. Plink v1.9 was used to identify Mendelian errors in the SCT sample set, for whom parental genotypes were available. For samples with Mendelian errors individual genotypes were removed. Genotype rates were 100% per SNP following imputation. Appendix 10 lists the number of SNPs analysed per gene for each sample set and test.

Tests of association with the three quantitative phenotypes: nonword repetition, language factor and global neurodevelopment index, were performed using Rvtests (Zhan *et al.*, 2016) in the SCT and Twin sample sets. For rare variants we used the Zeggini test (Morris & Zeggini, 2010) which is an aggregation test suitable for variants with $MAF < .01$. It computes a cumulative score of rare variant burden across the gene of interest which is then tested for association to a phenotype. The four X chromosome genes were excluded from the analysis for the SCT sample as genotypes did not follow conventional identification.

In a further analysis we used the SKAT test, which handles both common and rare variants within a single test (MAF upper limit 0.5). Both the Zeggini and the SKAT tests group the variants into a unit, in this case a gene.

Results

Figure 5 shows the distributions of scores on the three phenotypes for children with sex chromosome trisomies and the two comparison groups. The scores for Global burden are inverted so a low score corresponds to impairment, to be consistent with the other two measures, and scores on Global burden and Nonword repetition are jittered vertically as well as horizontally for clarity. Phenotypic characteristics of children with SCTs will be the focus of a separate publication, but we may note that, as anticipated, the group with SCTs show evidence of impairment on all three phenotype measures, but with a wide range of scores. For the combined sample with all cases ($N = 500$), nonword repetition correlated 0.76 with the language factor, and 0.60 with the global impairment rating. The language factor and global impairment rating correlated 0.69.

GSCA path-fitting analysis

Data from the three phenotypes were fitted using the model in Figure 4, first for the SCT group, and then separately for the two comparison groups. Table 2 shows the p-values based on 5000 permutations. The association with the 'Neurodev' factor did not meet our criterion for significance for either *NRXN1* nor *CNTNAP2* (see Table 2).

Exploratory analyses

Exploratory analyses were conducted to rule out the possibility that our analytic approach may have missed significant associations with specific SNPs. Appendix 9 shows associations computed for all the individual SNPs entered into the analysis, as well as the regression slopes relating number of minor alleles to phenotype for the two SNPs in *CNTNAP2* that have

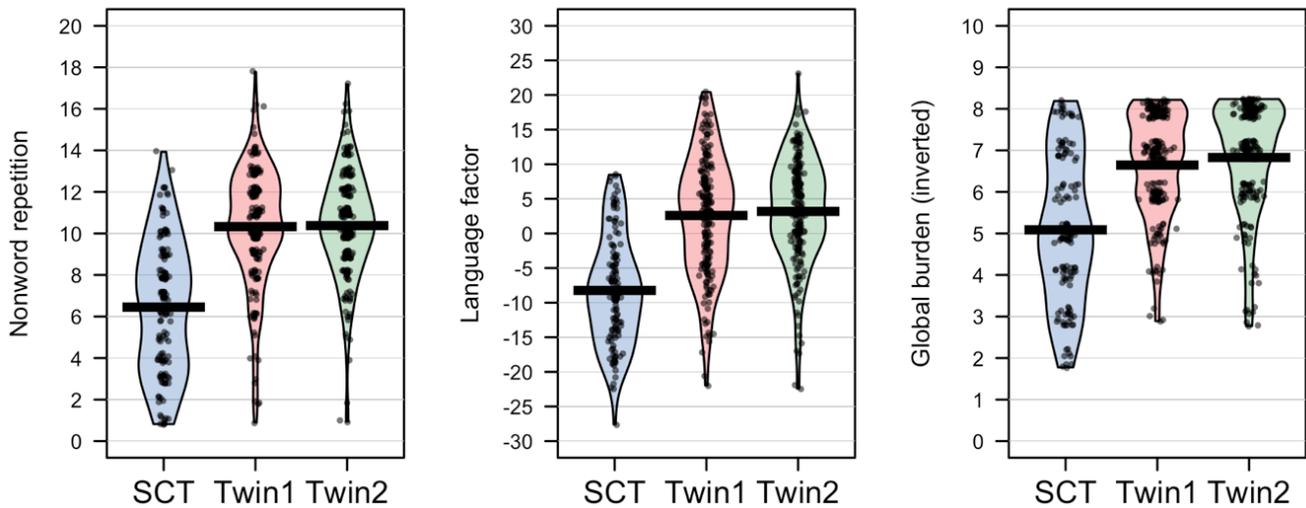


Figure 5. Pirate plots (Phillips, 2017) for sex chromosome trisomies (SCT) and two twin groups, showing individual cases as points, with bold line depicting median, for Nonword repetition (scaled score), Language factor and Global neurodevelopmental impairment (inverted so low score reflects impairment).

been a particular focus of research attention. The distribution of p-values seen for the whole collection of SNPs did not differ from that expected by chance.

In addition, we considered the suggestion by Skuse (2018) that results may vary by karyotype, because the expression of the X-linked homologue (*NLGN4X*) is variable, whereas in males with *XXY*, *NLGN4Y* is fully expressed. Although we did not assess *NLGN4X* and *NLGN4Y* directly, one might expect a more severe impact of a double hit in the *XXY* group, which would be reflected in the epistatic interactions with *NLGN*. To test this, we reran the GSCA separately for *XXX*, *XXY* and *XXY* subsamples, restricting consideration to those who were diagnosed prenatally, as specified in our protocol. As shown in Table 3, there was no hint of any association between genotype and phenotype in these subgroups. Note, however, as we originally stated, this analysis has very low power to detect true effects.

Further exploratory analyses added in version 3

Gene-based analyses. After Bonferroni correction for multiple testing none of the p-values were significant across the 22 candidate genes tested (Appendix 10) (Bishop *et al.*, 2019).

Discussion and conclusions

We assessed a “double-hit” model of differential susceptibility in children with sex-chromosome trisomies (SCTs), focusing on the hypothesis that the effects of common variations in neuroligin/neurexin genes upon language development are amplified in children with a sex chromosome trisomy (Bishop *et al.*, 2011). This hypothesis was evaluated through a targeted

Table 2. P-values from 5000 permutations for association with neurodevelopmental factor for *NRXN1* and *CNTNAP2*: results from Generalized Structured Component Analysis (GSCA).

Group	<i>NRXN1</i>	<i>CNTNAP2</i>
SCT	0.703	0.268
Twin1	0.080	0.323
Twin2	0.162	0.524

Table 3. P-values based on 5000 permutations for association with neurodevelopmental factor for *NRXN1* and *CNTNAP2* for sex chromosome trisomies (SCT) cases subdivided by karyotype, excluding those identified because of neurodevelopmental problems: results from Generalized Structured Component Analysis (GSCA).

Group	<i>NRXN1</i>	<i>CNTNAP2</i>
<i>XXX</i> , N = 31	0.311	0.256
<i>XXY</i> , N = 25	0.341	0.076
<i>XXY</i> , N = 23	0.787	0.447

analysis of two genes in the neuroligin/neurexin pathway, *CNTNAP2* and *NRXN1*, both of which have previously been associated with language development. A set of 47 SNPs were divided into corresponding candidate genes forming latent factors (weighted sums, in this case). These were analysed in relation to neurodevelopmental outcomes in a cohort of children with sex chromosome trisomies and two comparison twin groups, within a generalised structured component model (Hwang & Takane, 2004). Outcomes were represented by three related latent phenotype factors; non-word repetition, a measure of phonological short-term memory that has previously been associated with genetic variants associated with language and/or literacy, a general language factor score, and a measure of global neurodevelopmental impairment.

Children with SCTs showed evidence of impairment across all three of these phenotypes, indicating that they represent sensitive markers of the difficulties encountered by children with SCTs, and there was a wide range of severity of impairment. However, a factor derived from the three phenotypes was not associated with either candidate gene within the SCT sample or the two comparison datasets. Permutation testing confirmed that no effects reached our specified level of significance.

To address a concern that our method may mask associations with specific SNPs, we conducted an exploratory investigation of the individual variants within *NRXN1* and *CNTNAP2* (see Figure 4) using a more conventional regression approach testing for association with each of the three outcome factors (Appendix 9). The distribution of p-values across SNPs within each of the groups, aligned with that expected by chance, taking into account the correlations within the SNP set and the correlations between phenotypes.

Figure 6 shows results for two variants in *CNTNAP2* that had previously been robustly associated with language (rs779475 and rs207102). Previous studies indicate that for both of these SNPs the minor allele (T allele for rs779475 and G allele for rs207102) was associated with poorer neurodevelopmental outcomes (Alarcón *et al.*, 2008; Vernes *et al.*, 2008; Whitehouse *et al.*, 2011). However, in the current study, the direction of effects fluctuated between sample groups and outcome measures for both SNPs (see also Appendix 9) and did not approach statistical significance.

Lack of association

Considering first the failure to replicate associations between specific SNPs and phenotypes in the twin comparison samples, we should not be surprised by this result. Non-replication is common in genetic association studies and primarily reflects the complexity of the underlying genetic effects (Hirschhorn *et al.*, 2002; Ott, 2004). The majority of genetic contributions to language development are reported to have small effect sizes and are characteristically heterogeneous (Newbury *et al.*, 2014). Our sample size gave limited power to detect the kinds of effect size obtained in previous studies.

For the SCT sample, we argued that effect sizes would be larger if the double hit hypothesis was correct. However, it was still the case that unless we took steps to maximise the chances of association and to guard against type I errors we were at risk of obtaining spurious results (Grabitz *et al.*, 2018). This meant that we had to decide in advance which genes to focus on: it is entirely possible that a different selection of SNPs might have provided evidence for the double hit hypothesis: the challenge for this study was how to determine the optimal strategy for analysis. We used a structured literature review to

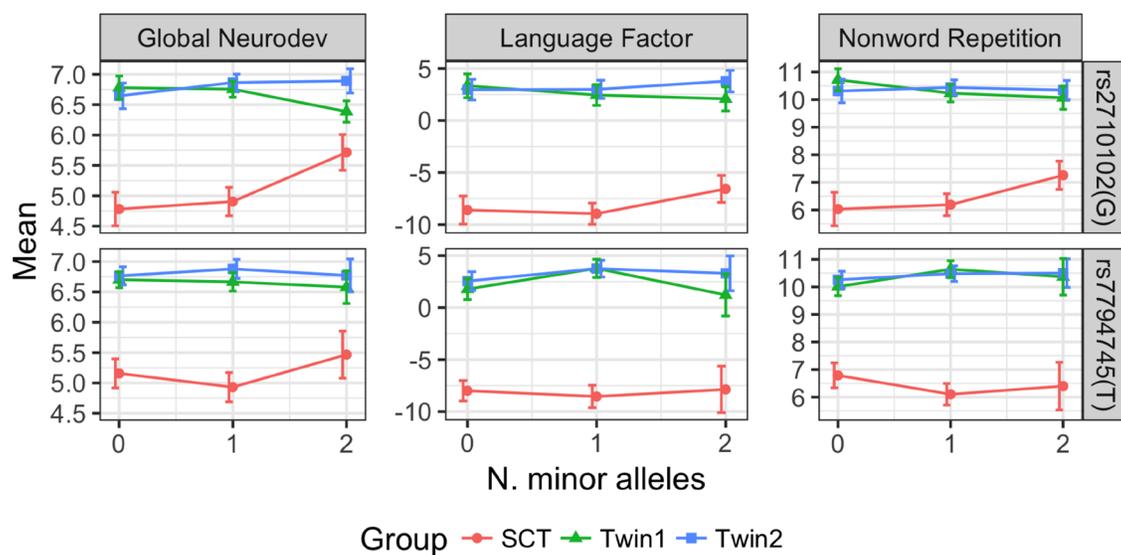


Figure 6. Mean scores for three phenotypes for SCT and both twin groups, in relation to number of minor alleles for rs2710102(G) and rs7794745(T).

identify genes implicated in developmental language disorder, generating a list of 41 candidate genes. Publication data for these genes were further scrutinised to narrow our search to a small number of genes with consistent and robust evidence for association. In this stage, we focused upon functional and behavioural dimensions of synaptic function and speech, language and communication respectively. Two genes scored highly on both of these dimensions and were selected for further literature-based review. One hundred and seven papers referring to *CNTNAP2* were examined. Of these, 32 included an analysis of common genetic variants and 19 examined rare variants of putative function. These papers tended to focus upon autism spectrum disorders but also considered speech and language and a diverse range of neurodevelopmental outcomes including intellectual disability and schizophrenia. Of the 32 common variant papers, 27 (84%) reported positive association, primarily with SNPs rs2710102 and/or rs7794745. The *CNTNAP2* gene has been implicated by both genome-wide screening approaches and targeted replication investigations and is reported to have a moderate effect size upon neurodevelopmental outcomes (Vernes *et al.*, 2008). One hundred and thirty one *NRXN1* papers were examined, including 45 reporting rare variants and 20 investigating common variants. For the *NRXN1* gene, there was a clear focus upon the relevance of copy number variants (deletions and duplications) in autistic disorder, epilepsy and schizophrenia in the rare variant studies while investigations of common variants focused upon nicotine dependency and working memory. Interestingly, a recent analysis of clinical microarray data showed that deletions in exons near the 5' end of *NRXN1* were specifically implicated in neurodevelopmental disorders (Lowther *et al.*, 2017). Given the large size of these two genes, we chose to focus our analysis on regions that have previously been associated with neurodevelopmental disorder, namely exons 13-14 in *CNTNAP2* and the 5' region of *NRXN1*. This evidence-based strategy reduced the possibility of false positives by providing a clear targeted rationale. However, it meant that our analyses were restricted to a small number of genetic variants within only two candidate genes. Within a complex genetic model, it is unlikely that one or two genes will explain a large proportion of the variance observed in language and communication, even under a hypothesis of increased susceptibility. Furthermore, our study design means that earlier identified candidates received more focus as they had a greater volume of supporting publications. We acknowledge that it is possible that we would have found association to alternative genes and/or SNPs if our search space had been extended but the chances of a type I error would also have increased (Grabitz *et al.*, 2018). Further exploratory gene-based analyses of 22 genes, as presented in the updated sections, indicate that this search-space would have to be considerably increased to yield positive associations. As such, the publication of a pre-registered strategy helped us to consider effect sizes and analysis strategies upfront, balancing out these considerations before the analysis stages.

Analysis methods

Genetic association studies typically consider each SNP as an independent variable giving rise to a particularly high number of tests. The typical significance threshold for a genome wide association study (GWAs) is 5×10^{-8} and a Bonferroni correction

for the 47 SNPs and three factors tested in the current study stands at 0.00035. Such adjustments are largely considered to be overly conservative as they do not account for the relationships between variants or the increased density of available genetic information (Fadista *et al.*, 2016). Recent developments have therefore been driven by gene-based analyses in which all variants within a gene are collapsed into a single factor for analysis. Gene-based methods reduce the number of tests required and can be applied to common and rare variants and allow for heterogeneous effects across a given gene (Lee *et al.*, 2014). Many gene-based methods exist, but most make quite specific assumptions and different tests can provide optimal power depending on study design (Lee *et al.*, 2014).

In this paper, we employed a gene-based analysis method which adapts a generalized structured component-based path model (GSCA) (Hwang & Takane, 2004) for genetic analysis. Previous studies indicate that this method is sensitive within relatively small sample sizes and enables the consideration of multiple related variables within a small number of latent factors (Romdhani *et al.*, 2015). Our own simulations (Appendix 1) indicated that this method was appropriate for a relatively small number of SNPs (<50) and traits, making it an attractive approach for a targeted hypothesis-led approach. Nonetheless, GSCA is a relatively novel approach to genetic analyses and the allowable number of variants remains small in genetic terms. More than 100 SNPs would be required to capture full information regarding common genetic variation across *CNTNAP2* and/or *NRXN1* and this dataset would be too large to give reliable results with our sample size.

Other gene-based tests based upon alternative statistical methods (e.g. burden tests, variance components and weighted analyses) are available and were applied to a limited number of candidate genes in our updated analyses. Even though these further exploratory investigations allowed a more comprehensive screening of genetic variation than the original analyses, which considered the effects of single SNPs, no significant associations were observed. These findings indicate that a much larger number of genes would need to be screened, which is not possible within the current dataset. As a general rule, the number of participants should be significantly greater than the number of parameters measured. Gene-based methods collapse effects across SNPs reducing the number of parameters to be estimated; however, adequate sample size should be ensured and not based on reducing dimensionality of data to achieve satisfactory power.

Alternative models of effects

Our strategy focused upon common genetic variants within a small number of well-supported candidate genes but it is clear that language and communication are complex traits which involve many interacting loci with different functional effects. Many candidate genes have been identified and characterised in the context of severe neurodevelopmental disorders and large hit events (rare variants and/or copy number changes) and it is possible that the effects of a given risk variant may be affected by genetic background, the mechanism of mutation and environmental factors (Weiner *et al.*, 2017). There are many possible mechanisms of functional interaction that would not be

detected by the approach taken in this study. For example, variants which affect gene expression (eQTLs) typically occur outside of coding regions and can have long-range effects (GTEx Consortium *et al.*, 2017). Since this study targeted restricted gene regions, we would not have detected long-range or trans- effects upon gene expression. Indeed, the control of gene expression is a complex process which involves many factors and this is complicated in this case by the involvement of X-inactivation and aneuploidy. Thus it would be of interest to directly measure the expression of the *NLGN4* genes and relate these to neurodevelopmental outcomes. Although this approach would only offer a snapshot of gene expression in time and space, it would be one step closer to contributory mechanisms and would provide a picture of the effects of aneuploidy upon gene expression and vice versa. Similarly, while a SNP-based analysis may tag structural rearrangements (deletions and duplications), an alternative approach would be required to fully assess the role of copy number variants. Copy number variants have been reported to play a role in neurodevelopmental disorder (Girirajan *et al.*, 2011) and developmental language disorders (Simpson *et al.*, 2015).

In summary, we did not observe association to either the *CNTNAP2* or *NRXN1* genes in our SCT cases or comparison controls, nor to a further 22 candidate genes in an extended analysis (version 3). We cannot reject the double-hit hypothesis on the basis of these analyses alone, but we can conclude that variants in the specific gene regions that we focused on do not appear to explain the phenotypic variation in neurodevelopment that is seen in children with an additional sex chromosome. In order to further the double hit hypothesis, additional analyses that consider many more candidate genes and/or different functional mechanisms will be required. Any such analyses would require larger sample sizes than was possible here and should be carefully controlled and pre-registered to avoid “fishing” experiments.

Self-certification statement

The authors confirm that they had no prior access to the full dataset described here.

For the Stage 1 Registered Report, the neurodevelopmental data had already been processed by DB and PT to derive the phenotypes to be used in the analysis; DFN had processed the DNA data separately to decide on the genotypes. The key tests proposed in this protocol involved putting the two strands of data together, which was deliberately not done, so that predictions could be derived without being aware of the data.

Data availability

Phenotype data, full tables of covariances and means of all variables, and an R markdown script for generating results are available on Open Science Framework. Dataset 1: SCT genetics: Double hit hypothesis <https://doi.org/10.17605/osf.io/pr54g>

Supplementary material

Appendices 1– 9 are available on [Open Science Framework](https://osf.io/u7dsw) (Bishop *et al.*, 2019). Appendix 10, containing additional analyses added to version 3, are also available at this web address.

Exploratory work: Appendix 10, which documents additional gene-based analyses that were not part of the pre-registered plan, is also available on Open Science Framework. <https://osf.io/u7dsw> (Bishop *et al.*, 2019)

The data is available under a CC0 1.0 Universal license.

Ethics approval for open data is conditional on data being fully anonymised. The genotype data could be used to identify individuals and so this part of the dataset cannot be made openly available. To optimise use of the data by others, we have created and deposited synthetic data using the R package *synthpop* v1.4-3 (Nowok *et al.*, 2016), which allows other users to explore a dataset that is closely based on the original data.

Acknowledgements

We offer warmest thanks to the families who took part in the study, and school staff who helped facilitate assessment arrangements. We greatly appreciate the help we had with recruitment from Unique (the Rare Chromosome Disorder Support Group), the Klinefelter’s Syndrome Association, and staff from Clinical Genetics Centres in the following NHS Trusts: Royal Devon and Exeter, The Northwest London Hospitals, University Hospitals Bristol, Sheffield Children’s, Liverpool Women’s, University Hospital Southampton, Birmingham Women’s Hospital, Great Ormond Street Hospital for Children, Leeds Teaching Hospitals, Newcastle upon Tyne Hospitals, Oxford University Hospitals, Central Manchester University Hospitals, University Hospitals of Leicester, Nottingham University Hospitals, Belfast Health and Social Care Trust and the NIHR Clinical Research Network. We extend special thanks to Dr Gillian Borthwick, National Research Coordinator NIHR CRN: Genetics, for helping us navigate the complexities of working with multiple NHS sites under the NIHR UK Rare Genetic Disease Research Consortium Agreement. The study would not have been possible without the generous support of Prof Simon Fisher, Director of the Max Planck Institute for Psycholinguistics, who offered facilities for analysis of genetic material during the data collection phase of this study. We thank Katerina Kucera and Arianna Vino for their substantial work in processing DNA samples.

We thank the High-Throughput Genomics Group at the Wellcome Trust Centre for Human Genetics (funded by Wellcome Trust grant reference 090532/Z/09/Z) for the generation of the Sequencing data.

Finally, the study would not have been possible without the hard work and dedication of eight research assistants who travelled around the UK to conduct assessments, including Nicola Gratton, Georgina Holt, Annie Brookman, Elaine Gray, Louise Atkins, Holly Thornton, Sarah Morris and Eleanor Paine.

References

- Abrahams BS, Geschwind DH: **Connecting genes to brain in the autism spectrum disorders.** *Arch Neurol.* 2010; **67**(4): 395–399.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Alarcón M, Abrahams BS, Stone JL, *et al.*: **Linkage, association, and gene-expression analyses identify *CNTNAP2* as an autism-susceptibility gene.** *Am J Hum Genet.* 2008; **82**(1): 150–159.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- American Psychiatric Association: **Diagnostic and statistical manual of mental disorders, Fifth Edition.** Arlington, VA: American Psychiatric Publishing. 2013.
[Publisher Full Text](#)
- Anderson CA, Pettersson FH, Clarke GM, *et al.*: **Data quality control in genetic case-control association studies.** *Nat Protoc.* 2010; **5**(9): 1564–1573.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Arking DE, Cutler DJ, Brune CW, *et al.*: **A common genetic variant in the neurexin superfamily member *CNTNAP2* increases familial risk of autism.** *Am J Hum Genet.* 2008; **82**(1): 160–164.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Berlitch JB, Yang F, Xu J, *et al.*: **Genes that escape from X inactivation.** *Hum Genet.* 2011; **130**(2): 237–245.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Bishop DV: **The Children's Communication Checklist, version 2 (CCC-2).** London: Pearson. 2003.
[Reference Source](#)
- Bishop DV: **What causes specific language impairment in children?** *Curr Dir Psychol Sci.* 2006; **15**(5): 217–221.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Bishop DV: **Specific language impairment, dyslexia, and autism: Using genetics to unravel their relationship.** In CF Norbury, JB Tomblin, & DVM Bishop (Eds.), *Understanding developmental language disorders: from theory to practice.* Hove: Psychology Press, 2008; 67–78.
[Reference Source](#)
- Bishop DV: **Overlaps between autism and language impairment: phenomimicry or shared etiology?** *Behav Genet.* 2010; **40**(5): 618–629.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Bishop DV, Jacobs PA, Lachlan K, *et al.*: **Autism, language and communication in children with sex chromosome trisomies.** *Arch Dis Child.* 2011; **96**(10): 954–959.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Bishop DV, North T, Donlan C: **Nonword repetition as a behavioural marker for inherited language impairment: evidence from a twin study.** *J Child Psychol Psychiatry.* 1996; **37**(4): 391–403.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Bishop DV, Scerif G: **Klinefelter syndrome as a window on the aetiology of language and communication impairments in children: the neurologigin-neurexin hypothesis.** *Acta Paediatr.* 2011; **100**(6): 903–907.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Bishop DVM, Thompson PA, Newbury D, *et al.*: **SCT genetics: Double hit hypothesis.** 2019.
[Reference Source](#)
- Blasi F, Bacchelli E, Pesaresi G, *et al.*: **Absence of coding mutations in the X-linked genes *neurexin 3* and *neurexin 4* in individuals with autism from the IMGSAC collection.** *Am J Med Genet B Neuropsychiatr Genet.* 2006; **141B**(3): 220–221.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Cao X, Tabuchi K: **Functions of synapse adhesion molecules neurexin/neurologins and neurodevelopmental disorders.** *Neurosci Res.* 2017; **116**: 3–9.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Carrel L, Brown CJ: **When the Lyon(ized chromosome) roars: ongoing expression from an inactive X chromosome.** *Philos Trans R Soc Lond B Biol Sci.* 2017; **372**(1733): pii: 20160355.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Carrion-Castillo A, Maassen B, Franke B, *et al.*: **Association analysis of dyslexia candidate genes in a Dutch longitudinal sample.** *Eur J Hum Genet.* 2017; **25**(4): 452–460.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Chin WW, Newstead PR: **Structural equation modeling analysis with small samples using partial least squares.** In: Hoyle RH, (Ed.), *Statistical strategies for small sample research.* Thousand Oaks: Sage Publications. 1999; 307–341.
[Reference Source](#)
- Ching MS, Shen Y, Tan WH, *et al.*: **Deletions of *NRXN1* (neurexin-1) predispose to a wide spectrum of developmental disorders.** *Am J Med Genet B Neuropsychiatr Genet.* 2010; **153B**(4): 937–947.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Chocholska S, Rossier E, Barbi G, *et al.*: **Molecular cytogenetic analysis of a familial interstitial deletion Xp22.2-22.3 with a highly variable phenotype in female carriers.** *Am J Med Genet A.* 2006; **140**(6): 604–610.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Constantino JN: **The Social Responsiveness Scale.** Los Angeles: Western Psychological Services. 2005.
[Reference Source](#)
- Craig AM, Kang Y: **Neurexin-neurologin signaling in synapse development.** *Curr Opin Neurobiol.* 2007; **17**(1): 43–52.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Das S, Forer L, Schönherr S, *et al.*: **Next-generation genotype imputation service and methods.** *Nat Genet.* 2016; **48**(10): 1284–1287.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Fadista J, Manning AK, Florez JC, *et al.*: **The (in)famous GWAS P-value threshold revisited and updated for low-frequency variants.** *Eur J Hum Genet.* 2016; **24**(8): 1202–5.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Frederickson N, Frith U, Reason R: **Phonological assessment battery (PhAB).** Windsor: NFER-Nelson. 1997.
[Reference Source](#)
- Gauthier J, Bonnel A, St-Onge J, *et al.*: ***NLGN3/NLGN4* gene mutations are not responsible for autism in the Quebec population.** *Am J Med Genet B Neuropsychiatr Genet.* 2005; **132B**(1): 74–75.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Girirajan S, Brkanac Z, Coe BP, *et al.*: **Relative burden of large CNVs on a range of neurodevelopmental phenotypes.** *PLoS Genet.* 2011; **7**(11): e1002334.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Girirajan S, Rosenfeld JA, Cooper GM, *et al.*: **A recurrent 16p12.1 microdeletion supports a two-hit model for severe developmental delay.** *Nat Genet.* 2010; **42**(3): 203–209.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Goodman R, Ford T, Richards H, *et al.*: **The Development and Well-Being Assessment: description and initial validation of an integrated assessment of child and adolescent psychopathology.** *J Child Psychol Psychiatry.* 2000; **41**(5): 645–655.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Grabitz CR, Button KS, Munafò MR, *et al.*: **Logical and methodological issues affecting genetic studies of humans reported in top neuroscience journals.** *J Cogn Neurosci.* 2018; **30**(1): 25–41.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Graham SA, Fisher SE: **Understanding Language from a Genomic Perspective.** *Ann Rev Genet.* 2015; **49**: 131–160.
[PubMed Abstract](#) | [Publisher Full Text](#)
- GTEx Consortium, Laboratory, Data Analysis & Coordinating Center (LDACC)—Analysis Working Group, Statistical Methods groups—Analysis Working Group, *et al.*: **Genetic effects on gene expression across human tissues.** *Nature.* 2017; **550**(7675): 204–213.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Harris PA, Taylor R, Thielke R, *et al.*: **Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support.** *J Biomed Inform.* 2009; **42**(2): 377–381.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Hirschhorn JN, Lohmueller K, Byrne E, *et al.*: **A comprehensive review of genetic association studies.** *Genet Med.* 2002; **4**(2): 45–61.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Hussain NK, Sheng M: **Neuroscience. Making synapses: a balancing act.** *Science.* 2005; **307**(5713): 1207–1208.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Hwang H, Takane Y: **Generalized structured component analysis.** *Psychometrika.* 2004; **69**(1): 81–99.
[Publisher Full Text](#)
- Jamain S, Quach H, Betancur C, *et al.*: **Mutations of the X-linked genes encoding neurologins *NLGN3* and *NLGN4* are associated with autism.** *Nat Genet.* 2003; **34**(1): 27–29.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Jaramillo TC, Escamilla CO, Liu S, *et al.*: **Genetic background effects in *Neurologin-3* mutant mice: Minimal behavioral abnormalities on C57 background.** *Autism Res.* 2018; **11**(2): 234–244.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Korkman M, Kirk U, Kemp SI: **NEPSY: A developmental neuropsychological assessment.** San Antonio: Psychological Corporation. 1998.
[Reference Source](#)
- Laumonnier F, Bonnet-Brilhault F, Gomot M, *et al.*: **X-linked mental retardation and autism are associated with a mutation in the *NLGN4* gene, a member of the neurologin family.** *Am J Med Genet.* 2004; **74**(3): 552–557.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Lawson-Yuen A, Saldivar JS, Sommer S, *et al.*: **Familial deletion within *NLGN4* associated with autism and Tourette syndrome.** *Eur J Hum Genet.* 2008; **16**(5): 614–618.
[PubMed Abstract](#) | [Publisher Full Text](#)

- Lee S, Abecasis GR, Boehnke M, et al.: **Rare-variant association analysis: study designs and statistical tests.** *Am J Hum Genet.* 2014; **95**(1): 5–23.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Leggett V, Jacobs P, Nation K, et al.: **Neurocognitive outcomes of individuals with a sex chromosome trisomy: XXX, XYY, or XXY: a systematic review.** *Dev Med Child Neurol.* 2010; **52**(2): 119–129.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Linden MG, Bender BG: **Fifty-one prenatally diagnosed children and adolescents with sex chromosome abnormalities.** *Am J Med Genet.* 2002; **110**(1): 11–18.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Liu Y, Du Y, Liu W, et al.: **Lack of association between NLGN3, NLGN4, SHANK2 and SHANK3 gene variants and autism spectrum disorder in a Chinese population.** *PLoS One.* 2013; **8**(2): e56639.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Lowther C, Speevak M, Armour CM, et al.: **Molecular characterization of NRXN1 deletions from 19,263 clinical microarray cases identifies exons important for neurodevelopmental disease expression.** *Genet Med.* 2017; **19**(1): 53–61.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Lu Z, Reddy MV, Liu J, et al.: **Molecular Architecture of Contactin-associated Protein-like 2 (CNTNAP2) and Its Interaction with Contactin 2 (CNTN2).** *J Biol Chem.* 2016; **291**(46): 24133–24147.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Macarow M, Zeigler M, Newman JP, et al.: **Deletions of VCX-A and NLGN4: a variable phenotype including normal intellect.** *J Intellect Disabil Res.* 2007; **51**(Pt 5): 329–333.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Marino C, Meng H, Mascheretti S, et al.: **DCDC2 genetic variants and susceptibility to developmental dyslexia.** *Psychiatr Genet.* 2012; **22**(1): 25–30.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Marshall CR, Noor A, Vincent JB, et al.: **Structural variation of chromosomes in autism spectrum disorder.** *Am J Hum Genet.* 2008; **82**(2): 477–488.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Morris AP, Zeggini E: **An evaluation of statistical approaches to rare variant analysis in genetic association studies.** *Genet Epidemiol.* 2010; **34**(2): 188–93.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Neale MD: **Neale Analysis of Reading Ability Second Revised British Edition.** London: GL Assessment. 1999.
[Reference Source](#)
- Newbury DF, Mari F, Sadighi Akha E, et al.: **Dual copy number variants involving 16p11 and 6q22 in a case of childhood apraxia of speech and pervasive developmental disorder.** *Eur J Hum Genet.* 2013; **21**(4): 361–365.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Newbury DF, Monaco AP, Paracchini S: **Reading and language disorders: the importance of both quantity and quality.** *Genes (Basel).* 2014; **5**(2): 285–309.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Newbury DF, Paracchini S, Scerri TS, et al.: **Investigation of dyslexia and SLI risk variants in reading- and language-impaired subjects.** *Behav Genet.* 2011; **41**(1): 90–104.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Newbury DF, Simpson NH, Thompson PA, et al.: **Stage 1 Registered Report: Variation in neurodevelopmental outcomes in children with sex chromosome trisomies: protocol for a test of the double hit hypothesis [version 2; referees: 3 approved].** *Wellcome Open Res.* 2018; **3**: 10.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Newbury DF, Winchester L, Addis L, et al.: **CMIP and ATP2C2 modulate phonological short-term memory in language impairment.** *Am J Hum Genet.* 2009; **85**(2): 264–272.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Nielsen J, Wohler M: **Chromosome abnormalities found among 34,910 newborn children: results from a 13-year incidence study in Arhus, Denmark.** *Hum Genet.* 1991; **87**(1): 81–83.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Norbury CF, Gooch D, Wray C, et al.: **The impact of nonverbal ability on prevalence and clinical presentation of language disorder: evidence from a population study.** *J Child Psychol Psychiatry.* 2016; **57**(11): 1247–1257.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Nowok B, Raab GM, Dibben C: **synthpop: Bespoke Creation of Synthetic Data in R.** *J Stat Softw.* 2016; **74**(11): 1–26.
[Publisher Full Text](#)
- Ott J: **Association of genetic loci: replication or not, that is the question.** *Neurology.* 2004; **63**(6): 995–8.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Pampanos A, Volaki K, Kanavakis E, et al.: **A substitution involving the NLGN4 gene associated with autistic behavior in the Greek population.** *Genet Test Mol Biomarkers.* 2009; **13**(5): 611–5.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Phillips N: **yarr: A Companion to the e-Book "YaRrr!: The Pirate's Guide to R".** R package version 0.1.5. 2017.
[Reference Source](#)
- Price AL, Patterson NJ, Plenge RM, et al.: **Principal components analysis corrects for stratification in genome-wide association studies.** *Nat Genet.* 2006; **38**(8): 904–909.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Purcell S, Neale B, Todd-Brown K, et al.: **PLINK: a tool set for whole-genome association and population-based linkage analyses.** *Am J Hum Genet.* 2007; **81**(3): 559–575.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- R Core Team: **R: A language and environment for statistical computing.** Vienna, Austria: R Foundation for Statistical Computing. 2016.
[Reference Source](#)
- Rice ML, Zubrick SR, Taylor CL, et al.: **Longitudinal Study of Language and Speech of Twins at 4 and 6 Years: Twinning Effects Decrease, Zygosity Effects Disappear, and Heritability Increases.** *J Speech Lang Hear Res.* 2018; **61**(1): 79–93.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Romdhani H, Grinek S, Hwang H, et al.: **ASGSCA: Association Studies for multiple SNPs and multiple traits using Generalized Structured Equation Models.** R package version 1.12.0. 2014.
[Reference Source](#)
- Romdhani H, Hwang H, Paradis G, et al.: **Pathway-based association study of multiple candidate genes and multiple traits using structural equation models.** *Genet Epidemiol.* 2015; **39**(2): 101–113.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Ross JL, Roeltgen DP, Kushner H, et al.: **Behavioral and social phenotypes in boys with 47,XXY syndrome or 47,XXY Klinefelter syndrome.** *Pediatrics.* 2012; **129**(4): 769–778.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Scerri TS, Morris AP, Buckingham LL, et al.: **DCDC2, KIAA0319 and CMIP are associated with reading-related traits.** *Biol Psychiatry.* 2011; **70**(3): 237–245.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Simmons JP, Nelson LD, Simonsohn U: **A 21 word solution.** 2012.
[Publisher Full Text](#)
- Simpson NH, Addis L, Brandler WM, et al.: **Increased prevalence of sex chromosome aneuploidies in specific language impairment and dyslexia.** *Dev Med Child Neurol.* 2014; **56**(4): 346–353.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Simpson NH, Ceroni F, Reader RH, et al.: **Genome-wide analysis identifies a role for common copy number variants in specific language impairment.** *Eur J Hum Genet.* 2015; **23**(10): 1370–7.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Skuse D: **Referee Report For: Stage 1 Registered Report: Variation in neurodevelopmental outcomes in children with sex chromosome trisomies: protocol for a test of the double hit hypothesis [version 1; referees: 1 approved, 2 approved with reservations].** *Wellcome Open Res.* 2018; **3**: 10.
[Publisher Full Text](#)
- Talebizadeh Z, Lam DY, Theodoro MF, et al.: **Novel splice isoforms for NLGN3 and NLGN4 with possible implications in autism.** *J Med Genet.* 2006; **43**(5): e21.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Tenenhaus M, Vinzi VE, Chatelin YM, et al.: **PLS path modeling.** *Comput Stat Data Anal.* 2005; **48**(1): 159–205.
[Publisher Full Text](#)
- Thompson PA, Bishop DVM, Eising E, et al.: **Generalized Structured Component Analysis in candidate gene association studies: applications and limitations [version 2; peer review: 3 approved].** *Wellcome Open Res.* 2020; **4**: 142.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Torgesen JK, Wagner R, Rashotte C: **Test of Word Reading Efficiency (TOWRE).** New York: Psychological Corporation. 1999.
[Reference Source](#)
- Vernes SC, Newbury DF, Abrahams BS, et al.: **A functional genetic link between distinct developmental language disorders.** *N Engl J Med.* 2008; **359**(22): 2337–2345.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Vincent JB, Kolozsvari D, Roberts WS, et al.: **Mutation screening of X-chromosomal neurologigin genes: no mutations in 196 autism probands.** *Am J Med Genet B Neuropsychiatr Genet.* 2004; **129B**(1): 82–84.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Wechsler D: **Wechsler Abbreviated Scale of Intelligence.** San Antonio: Psychological Corporation. 1999.
[Reference Source](#)
- Weiner DJ, Wigdor EM, Ripke S, et al.: **Polygenic transmission disequilibrium confirms that common and rare variation act additively to create risk for autism spectrum disorders.** *Nat Genet.* 2017; **49**(7): 978–85.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Whitehouse AJ, Bishop DV, Ang QW, et al.: **CNTNAP2 variants affect early language development in the general population.** *Genes Brain Behav.* 2011; **10**(4): 451–6.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Wilson AC, Bishop DVM: **Resounding failure to replicate links between developmental language disorder and cerebral lateralisation.** *PeerJ.* 2018a; **6**: e4217.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Wilson AC, Bishop DVM: **Sex chromosome trisomies are not associated with**

atypical lateralization for language. *Dev Med Child Neurol.* 2018b; **60**(11): 1123–1139.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Woodcock RW, McGrew KS, Mather N: **Woodcock-Johnson III Tests of Cognitive Abilities.** Rolling Meadows, IL Riverside Publishing. 2007.

Yan J, Feng J, Schroer R, *et al.*: **Analysis of the neuroligin 4Y gene in patients with autism.** *Psychiatr Genet.* 2008; **18**(4): 204–207.

[PubMed Abstract](#) | [Publisher Full Text](#)

Yan J, Oliveira G, Coutinho A, *et al.*: **Analysis of the neuroligin 3 and 4 genes in autism and other neuropsychiatric patients.** *Mol Psychiatry.* 2005; **10**(4): 329–332.

[PubMed Abstract](#) | [Publisher Full Text](#)

Yanagi K, Kaname T, Wakui K, *et al.*: **Identification of four novel synonymous**

substitutions in the x-linked genes *neuroligin 3* and *neuroligin 4X* in japanese patients with autistic spectrum disorder. *Autism Res Treat.* 2012; **2012**: 724072.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Zhan X, Hu Y, Li B, *et al.*: **RVTESTS: an efficient and comprehensive tool for rare variant association analysis using sequence data.** *Bioinformatics.* 2016; **32**(9): 1423–1426.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Zweier C, de Jong EK, Zweier M, *et al.*: ***CNTNAP2* and *NRXN1* are mutated in autosomal-recessive Pitt-Hopkins-like mental retardation and determine the level of a common synaptic protein in *Drosophila*.** *Am J Hum Genet.* 2009; **85**(5): 655–666.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Open Peer Review

Current Peer Review Status:  

Version 4

Reviewer Report 03 June 2021

<https://doi.org/10.21956/wellcomeopenres.18694.r44186>

© 2021 Pourcain B. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Beate St Pourcain 

Language and Genetics Department, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

I approve of the amended article.

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 3

Reviewer Report 04 January 2021

<https://doi.org/10.21956/wellcomeopenres.17228.r40439>

© 2021 Skuse D. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



David H. Skuse 

Behavioural Sciences Unit, UCL Great Ormond Street Institute of Child Health, London, UK

In this, the third version of their planned test of the 'double hit' hypothesis, Newbury *et al.* have now tested their hypothesis that the adverse impact of an additional sex chromosome on language development could be amplified by genes that are expressed from the sex chromosomes (specifically NLGN3 and NLGN4) in epistasis with specific autosomal genes NRXN1 and CNTNAP2. In my original commentary on the proposal (Version 1) I suggested that the

evidence supporting the hypothesis was rather weak. However, it is undeniable that there is a considerable variation in the phenotypic presentation of the sex chromosome aneuploidies (XXX, XXY, XYY), which exceeds variation in the general population in terms of their favoured measures of language competence (see Figure 5 in Version 3). We do not yet have any satisfactory explanation for why that should be the case.

In the current report, no association was found between phenotypic variation and specific potentially salient variants of either NRXN1 or CNTNAP2. Intriguingly, previous research that has linked a disordered language phenotype to specific minor allelic variation of CNTNAP2 was not replicated here in a typically developing twin comparison sample. This could be due to a lack of power to detect an effect, as the finding was previously replicated in many studies. Consequently, it seems reasonable to conclude the analysis that employed a novel statistical methodology (critiqued in my previous commentary), which was supposedly suitable for small sample sizes, was probably subject to a Type II error.

The authors went on to look at alternative models of effects in a small number of candidate genes that have previously been associated with language-related phenotypes, and hence could modify the impact of any epistatic interactions between their X-linked candidates (NLGN3/4) and NRXN1 or CNTNAP2 variants. There is an image describing how these genes were selected from a SCOPUS search provided in Figure S1. In Appendix 9 an analysis is described for 23 SNPs in NRXN1 and 24 in CNTNAP2 where association was sought with language phenotypes in both the SCA and the twin samples. After Bonferroni correction, no association met statistical significance. In the supplementary material, the SNP analyses of the NRXN1 gene for variants 7-11 were omitted, presumably in error.

Gene-based analyses were performed on the 22 genes described in Figure S1 (Appendix 5), which apparently involved 'all known variants across the 22 candidate genes' and their potential association with the three key language variables of interest. I am unfamiliar with this mode of analysis so cannot comment on the specific risk of false positive results. Little is said about the outcome of this analysis other than that, after Bonferroni correction, none of the p-values was 'significant' – meaning that (presumably) variations in these additional genes did not influence the strength of association that was the subject of the original hypothesis.

I posited in my original commentary that CNVs could be impacting on language development in their sample of sex chromosome aneuploidies and acting as vulnerability factors. I am pleased to see that the authors are planning to undertake these analyses in the future. They may also examine the variability in the 'escape from inactivation' status of NLGN4X, which lies outside the PAR1 and whose inactivation is believed to be only partial.

I also predicted that there would be least variability in the language-related phenotype of the XYY male sample and arguably the most variable phenotypes in the XXX sample, because of the relationship between NLGN4X (partially inactivated) and NLGN4Y (fully expressed). In this latest version of their paper the authors state they followed up my suggestion and provide the results of their investigation in Table 3. Table 3 as presented on p11 does not mention NLGN4 but illustrates a completely different analysis relating to NRXN1 and CNTNAP2. It is nevertheless stated that my hypothesis was not supported!

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Neurodevelopmental disorders, sex chromosomal aneuploidies, genetic syndromes associated with intellectual disability.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 20 May 2021

Dorothy Bishop, University of Oxford, Oxford, UK

Many thanks for these additional comments on our paper. We have made some modifications to the main text, and in addition offer some replies here to specific points:

1. Power of GSCA. Since writing our original paper, we have done further analysis of the power of GSCA using simulations (Thompson, P. A., Bishop, D. V. M., Eising, E., Fisher, S. E., & Newbury, D. F. (2020). Generalized Structured Component Analysis in candidate gene association studies: Applications and limitations. *Wellcome Open Research*, 4, 142. <https://doi.org/10.12688/wellcomeopenres.15396.2>). A type II error cannot be ruled out, but we have shown that the method works well with the kind of sample size and number of SNPs used here.

2. Omission of data from table. Thanks to the reviewer for spotting the omission of one set of SNPs in Appendix 9. We have corrected Appendix 9 (on Open Science Framework) to now include the missing SNPs, and have updated the script used to generate the tables.

3. Clarification regarding gene-based analyses. The gene-based analyses involve a cumulative assessment of all genetic variation in a single test allowing for differences in functional variants across individuals. The results indicate that variations across the 22 genes tested do not influence language outcomes in SCT cases or twins.

4. Role of CNVs.

Since publishing version 3, we have reported an analysis of CNVs in this sample. Once again we have null findings. See: Mountford, H. S., Bishop, D. V. M., Thompson, P. A., Simpson, N. H., & Newbury, D. F. (2020). Copy number variation burden does not predict severity of neurodevelopmental phenotype in children with a sex chromosome trisomy. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics*, 184(2), 256–266. <https://doi.org/10.1002/ajmg.c.31791>

5. Skuse predictions and Table 3.

Table 3 tests for putative epistatic effects of variations across *NRXN1* and *CNTNAP2* in SCT subgroups, in line with the analyses presented elsewhere in the manuscript. These analyses do not directly test *NLGN4* but instead focus upon epistatic interactions with neuroligins. The idea is that if there is more consistent expression of *NLGN4Y*, then this might be evident as stronger impact in epistatic interactions. We wanted to credit the reviewer for the idea of expression differences in *NLGN4YX* and *NLGN4Y*, but have now reworded to clarify that the analysis in Table 3 is indirect evidence for this effect.

Competing Interests: No competing interests were disclosed.

Reviewer Report 10 November 2020

<https://doi.org/10.21956/wellcomeopenres.17228.r40445>

© 2020 Pourcain B. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Beate St Pourcain 

Language and Genetics Department, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

Newbury and colleagues have extended their studies of variation in neurodevelopmental outcomes in children with sex chromosome trisomies with extensive gene-based analyses. However, none of the gene-based analyses passed the multiple testing threshold (Bonferroni correction) indicating that larger number of genes and samples will need to be screened to find evidence supporting the double-hit hypothesis.

I have only minor comments:

1. I assume the authors mean Indels and multi-allelic variants when referring to Non-SNPs. Could the authors please re-phrase?
2. Within the exploratory analyses paragraph, it reads as if imputation was used to impute also rare variants for rare-variant analyses. "Common and rare variants were further examined using the SKAT test (MAF upper limit 0.5)." Do the authors just mean a rare variant test applied to common variants? Please clarify which variants (MAF range) were analysed with which test.
3. Could the authors please clarify what they mean with the "Zeggini" test? "A rare variant test analysis was performed using Rvtests (Zhan et al, 2016) to analyse three quantitative phenotypes: nonword repetition, language factor and global neurodevelopment index in the SCT and Twin sample sets using the Zeggini test with a MAF upper limit of 0.01."
4. The author may consider increasing the number of digits reported for p (Appendix 9). For example, the TWA_global_p for rs2193417_G is given as 0.00 (NRXN1), which is now difficult to evaluate with respect to a multiple testing threshold of 0.0003.

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 19 Dec 2020

Dorothy Bishop, University of Oxford, Oxford, UK

We thank the reviewer for the additional points, which allow us to add further clarification.

1. Yes, the non-SNPs included indels and multi-allelic SNPs.

2-3. Two analyses were performed within RVTESTS; the SKAT test and the Zeggini test. The Zeggini test is a test of association. This is an aggregation test which computes a cumulative score of rare variant burden across the gene of interest which is then tested for association to the phenotype of interest. The test is described in doi: 10.1002/gepi.20450.

The Zeggini test is suited to the analysis of rare variants only, while SKAT can handle both common and rare variants within a single test. Accordingly, we applied the Zeggini test to all variants with $MAF \leq 0.01$ and the SKAT test to all variants with $MAF \leq 0.5$ (i.e. all variants across the gene, both common and rare).

4. Appendix 9 is not part of the additional content we added in version 3 of the manuscript - which is termed 'Exploratory analyses'. Appendix 9 was part of our original submission and was titled 'Additional exploratory analyses'. We apologise for confusing labelling of sections.

Competing Interests: No competing interests were disclosed.

Version 1

Reviewer Report 11 September 2018

<https://doi.org/10.21956/wellcomeopenres.15981.r33521>

© 2018 Skuse D. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



David H. Skuse 

Behavioural Sciences Unit, UCL Great Ormond Street Institute of Child Health, London, UK

The Stage 2 report summarizes the findings of the proposed investigation, which aimed to test the hypothesis that neuroligin genes on the sex chromosomes (specifically NLGN4X and Y) would be expressed in excessive dosage in multiple sex chromosome trisomy (SCT) aneuploidies (XXX,XXY,XXY) and that there could be an epistatic interaction between those genes and a specific autosomal gene on Chromosome 7 (CNTNAP2), variants of which have been shown to be associated with disorders of language in particular. Accordingly, this interaction would explain the deficits in language skills characterizing SCTs, which are relatively more common in XYY than XXX

individuals, with XXY males intermediate between the other two conditions.

In the original proposal, besides the epistatic interaction with CNTNAP2, the authors also proposed an interaction between NLGN and the autosomal gene NRXN1 (chromosome 2), such that variants in that gene within putative regulatory regions could modify the impact of excessive NLGN expression.

Essentially, the purpose of the investigation was to evaluate an intriguing 'double-hit' mechanism, whereby the variable impact of sex chromosome aneuploidy on language phenotypes would be explained by some sort of interaction with autosomal genetic variation, in two specific candidates (CNTNAP2 and NRXN1). The locus of the interaction was proposed to be at the synapse, reflected in variable synaptic scaffolding with functional consequences. Underlying this hypothesis was the observation that there is a transmembrane connection in the synaptic scaffolding system between CNTNAPs and Contactins/Catenins and also between Neuroligins/Neurexins as illustrated in Figure 1.

Having reviewed the material in the original Stage 1 Proposal, I suggested there should be the most variable language phenotypes within the XXX sample, and the least variable language phenotype in the XYY sample. I expressed skepticism that the overall double-hit hypothesis would be supported by the proposed analysis, and also that the novel statistical methods to be used would truly get around the problem of limited power (which arose from the small sample size). The Stage 2 report (19th July 2018) states there was wide phenotypic variation observed within the SCT groups, but no overall association could be found with the variants that had been measured on the autosomal genes CNTNAP2 or NXN1. Nor were there any specific associations within each of the trisomies investigated. The latest report reiterates much of the information about the study design that was provided by the Stage 1 report (and its April 2018 revision).

A variety of analyses are presented for which the overall structure is represented by Figure 4. It is gratifying that measures of language impairment, for which details will be given in a future publication, did show sensitivity to the presence of sex chromosome aneuploidy status, and good correlations with the language factor. I was not surprised that the findings reported in Figure 6, which attempted to replicate previous reports on associations between specific SNPs on CNTNAP2 and neurodevelopmental impairment, showed no statistically significant result. As the authors state, replication in these complex phenotype-genotype associations is the exception rather than the rule, and there is likely to be publication bias in favor of positive results. It is commendable that the authors did not continue to look for association with alternative genes/SNPs in a fishing expedition that would have led them open to the risk of a type 1 error.

There was no evidence found that the results varied by karyotype as I had proposed in my earlier commentary, on the grounds that there was likely to be a more variable language phenotype in XXX than XYY individuals because NLGN4X was subject to variable and partial X-inactivation (whereas NLGN4Y was fully expressed). Accordingly, all males with XYY would over-express NLGN4Y to the same extent, but females with XXX (in which two X-chromosomes are silenced) would express NLGN4X to a highly variable degree depending on individual differences in X-inactivation status.

I was surprised by the comment that, in the context of a discussion about alternative models of effects, that there are many possible mechanisms of functional interaction that would not be detected by the approach taken by this study – such as functionally valid variants that lie outside

the coding regions. My understanding is that many SNPs employed in this study were indeed outside coding regions (though may have had functional significance).

There is a further comment concerning a mechanism whereby there could be modulation of expression of NLGN4X/Y by autosomal variants: a point made in my original commentary is that the regulation of NLGN4X is complex in relation to X-inactivation, and that its expression in individuals with more than one X chromosome is unpredictable for that reason in particular. This mechanism leading to variability of expression may be more pertinent than modulation by autosomal genes.

Overall, the study was conducted with impeccable scientific rigor and exemplifies the benefits of registering a research design with Wellcome Open Research, and seeking peer review before undertaking the proposed analysis.

Are the data able to test the authors' proposed hypotheses by satisfying the approved outcome-neutral conditions (such as quality checks, positive controls)?

Yes

Are the introduction, rationale and stated hypotheses the same as the approved Stage 1 submission? (required)

Yes

Did the authors adhere precisely to the registered experimental procedures? If not, has an explanation been provided regarding any change?

Yes

Are any unregistered post hoc analyses added by the authors justified, methodologically sound and informative?

Yes

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

I cannot comment. A qualified statistician is required.

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Neurodevelopmental disorders, sex chromosomal aneuploidies, genetic syndromes associated with intellectual disability

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 31 August 2018

<https://doi.org/10.21956/wellcomeopenres.15981.r33522>

© 2018 Pourcain B. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Beate St Pourcain 

Language and Genetics Department, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

In this Stage 2 Registered Report Newbury and colleagues studied whether the presence of an extra sex chromosome is associated with an increased rate of neurodevelopmental difficulties involving language. The authors investigate here a double-hit hypothesis, specifically that the presence of an additional sex chromosome amplifies the impact of common autosomal genetic variants residing within genes involved in synaptic functionality. The authors specifically focus on the functional network of neuroligins and study common variants within *CNTNAP2* and *NRXN1*, and use complex analysis strategies, such as Generalised Structural Component Analysis, to overcome small sample size. Unfortunately, the sample size was small, and the authors found no support for the idea that common variants in *CNTNAP2* or *NRXN1* are associated with the severity of language and neurodevelopmental impairments that often accompany an extra X or Y chromosome.

However, although this study was small, the authors could make suggestions, based on their extensive expertise, how a future study should be conducted that may be able to overcome current limitations. It is highly commendable that authors followed a fully transparent analysis approach, sharing scripts and simulated comparable data. The authors may also want to highlight the transferability of their applied structural equation modelling approach.

Are the data able to test the authors' proposed hypotheses by satisfying the approved outcome-neutral conditions (such as quality checks, positive controls)?

Yes.

Are the introduction, rationale and stated hypotheses the same as the approved Stage 1 submission? (required)

Yes.

Did the authors adhere precisely to the registered experimental procedures? If not, has an explanation been provided regarding any change?

Yes.

Are any unregistered post hoc analyses added by the authors justified, methodologically sound and informative?

Yes.

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.
